DOCUMENT RESUME

ED 195 554 TM 800 579

TITLE Alaska Statewide Student Assessment: A Comparison of the 1977 and 1979 Assessment Results.
INSTITUTION Alaska State Dept. of Education, Juneau.
PUB DATE [80]
NOTE 29p.

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Academic Achievement: *Educational Assessment: Elementary Education: *Elementary School Mathematics: Grade 4: Grade 8: Outcomes of Education: *Reading Achievement: Scores: State Programs: *Testing Programs
IDENTIFIERS *Alaska: *Alaska Statewide Instructional Support System: State Assessment of Educational Progress

ABSTRACT

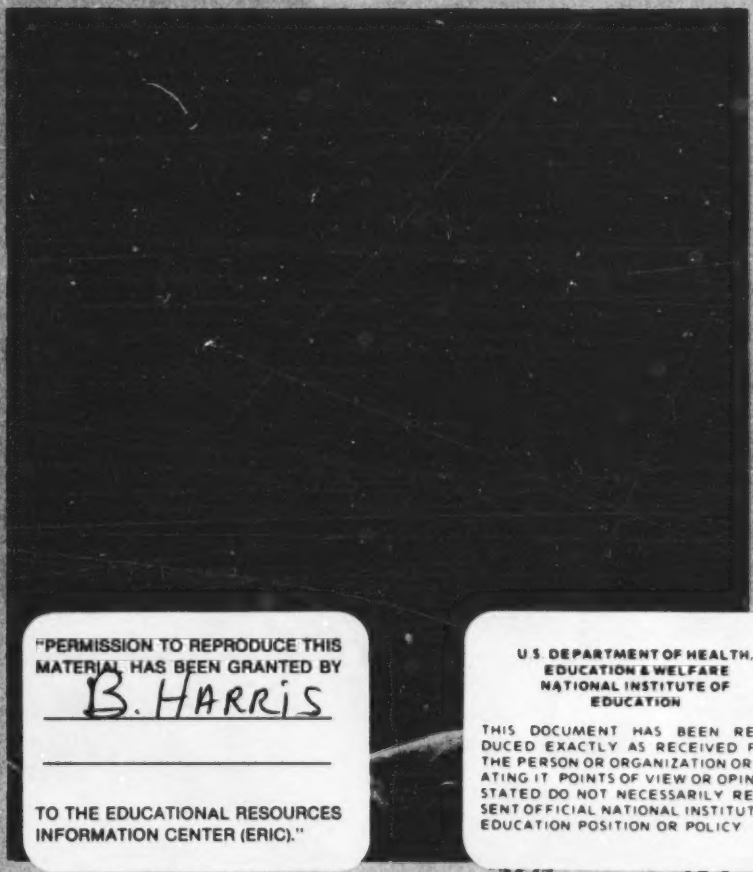
The results of the 1977 and 1979 Statewide Assessment Tests in Alaska are presented. This assessment testing is part of the testing component within the Alaska Statewide Instructional Support System (SISS), and is based on specific objectives critical to a student's academic success. Objectives covered are: mathematics computation, mathematics application, word identification and reading comprehension. The results are based on a statewide sample of students in grades 4 and 8. Student performance was compared to desired levels of performance on specific learning objectives. Statewide, scores did not change substantially between 1977 and 1979 for grade 4: across all content areas in 1977, students in small/rural districts scored lower than those in large or medium-sized districts. Bilingual students in grade 4 did better in 1979 than in 1977. Students in grade 8 showed similar patterns across all content areas: those from large and medium-sized districts performed better than those in small/rural districts for both years. Nonbilingual students performed better than bilingual students. In both grades, students with reading problems did not do as well as those without such problems. Statewide, 1979 performance results were better than the 1977 sample. (Author/GK)

****
* Reproductions supplied by EDRS are the b t that can be made *
* from the original document. *
****

SEP 1980
RECEIVED
ERIC/TM

Alaska Statewide Student Assessment:

A Comparison of the 1977 and 1979 Assessment Results



ED195554

TM 800579

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
B. HARRIS
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

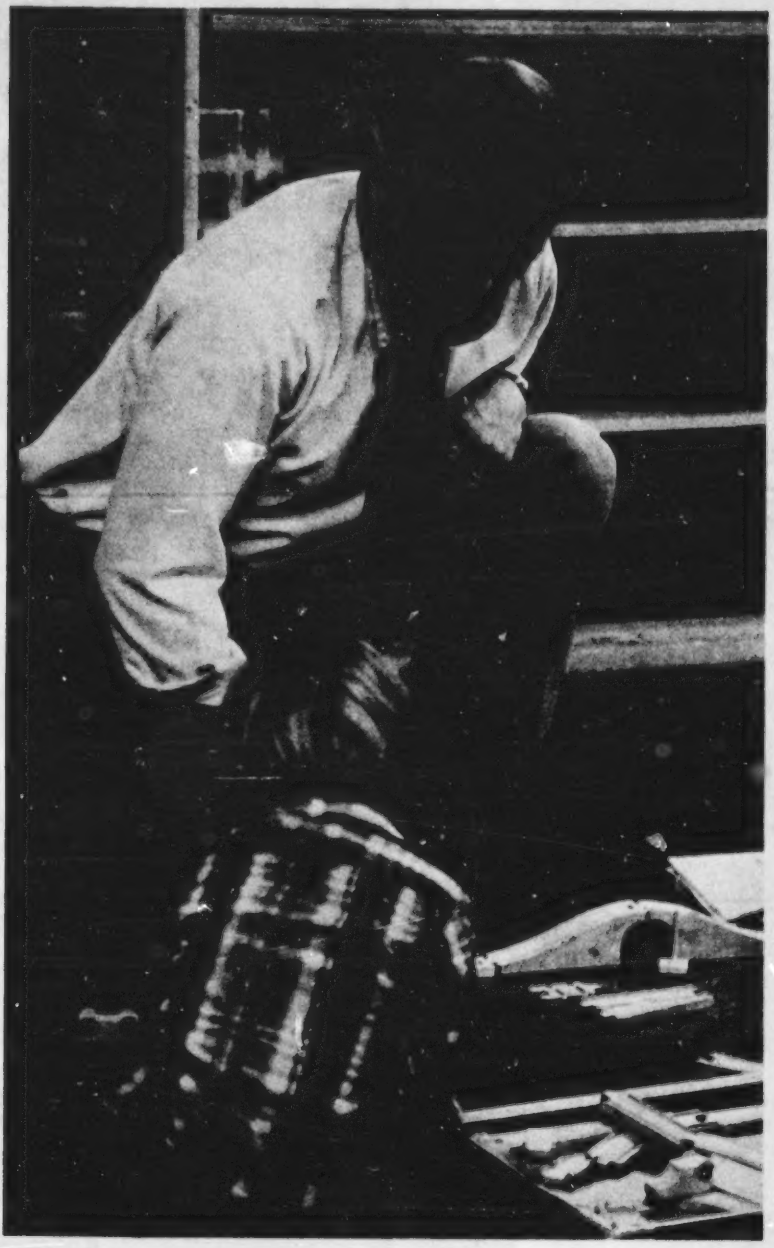
The Alaska Statewide Instructional Support

Alaska Statewide Student Assessment:

A Comparison of the 1977 and 1979 Assessment Results

The Alaska Statewide Instructional Support System

The earlier report of the 1977 statewide assessment contained inaccuracies due to a computer tape problem. This report replaces the former report.



FOREWORD

In the spring of 1977, the Alaska Department of Education initiated a biennial assessment of Alaskan elementary school children in reading and mathematics. All Alaskans can take pride in this program's history, for information provided through assessment will ultimately lead to improved learning opportunities for Alaska students.

At a 1974 meeting, educators from school districts around the state asked themselves: "How might we meet the public's desire to know how well Alaskan students are doing, and *at the same time* provide the professional educator with practical information for improving student skills?"

The answer to this two-part question is the Alaska Statewide Instructional Support System (SISS). The system contains four components:

1. **A Testing Component**
   To provide data about student learning at several levels.
2. **An Instructional Resource Component**
   To identify proven or promising instructional materials or practices related to the skills being measured.
3. **A Staff Development Component**
   To help teachers, administrators and students use the system to achieve their own educational goals.
4. **A Communications Component**
   To make the system accessible even in the most remote areas of Alaska.

The testing component is based on specific skills or objectives critical to a student's academic success. The tests are of two types: (1) diagnostic tests for individual and classroom use, and (2) statewide assessment tests. This report presents the results of the 1977 and 1979 Statewide Assessment Tests; the next statewide test will be administered in the spring of 1981. Technical information on the testing component is available from the Alaska State Department of Education.

The Alaska Statewide Instructional Support System is well underway. When successfully concluded, it will be one means of providing full educational opportunity to each Alaskan student. I personally want to thank those teachers, administrators and citizens who worked to make this possible. It is one more step in our mutual efforts to provide high quality public education.

Marshall L. Lind
Commissioner of Education

BACKGROUND and PROCEDURES

Who Was Tested?

The assessment results are based on a statewide sample of fourth and eighth grade students. The 1977 sample involved 112 schools, 2543 students; the 1979 sample, 114 schools and 2879 students. Samples were selected to represent a cross section of the state's student population by school size, race, and geographical region. Sampling provides accurate information efficiently and economically. When technically sound procedures are followed in selecting a sample, one can generalize from the sample to the larger population – in this case, all Alaska fourth and eighth graders.

A 20 percent classroom sample was drawn for each of the two grade levels. That is, each classroom selected for testing represented five classrooms of approximately the same size and student body. In the 1979 assessment several rural school districts failed to administer the assessment; this tends to distort the 1979 bilingual and rural comparisons.

The only students exempted from testing were those who were ill, absent or in a self-contained special education classroom. The statewide sample included "mainstreamed" students (special education students who study in the same classroom as regular students).

How and When Was the Test Administered?

The second week in April 1977 and the last week in March 1979 were designated as the dates for statewide assessment. In most sampled schools, teachers administered the test to their own students. Students provided certain biographical data (age, sex, race, etc.), while teachers provided more descriptive information such as whether the student was bilingual, whether a language other than English was spoken in the home, whether the student had a diagnosed reading problem, was participating in a compensatory education program, was receiving assistance in reading, or had a visual, hearing or speech impediment. Answers to these and similar questions enabled reviewers to see how well students with given characteristics had performed on the assessment.

What Were the Students Tested On?

Items in the statewide test are based on specific learning objectives. Two panels of teachers and curriculum specialists – one for math and one for reading – identified skills they felt should be acquired by the completion of grade 4 and grade 8. The panels applied high standards in making these judgments, considering only what they felt *should* be taught, regardless of whether these skills actually *were* being taught

in Alaska classrooms. Then, on the basis of these identified skills, they selected the objectives to be testing from a bank of objectives and items.

An objective is a goal stated in explicit, measurable terms: for example, "Students will be able to add two one-digit numerals correctly." Many test items could measure this objective (e.g., 1 + 2 = ? or 2 + 3 = ? or 9 + 7 = ?). Clearly, a test that included every potential item for every objective would be too long to be practical. In most cases, as with the Alaska Statewide Assessment, a few test items must be chosen to test each objective. The same objectives and items were used in the 1977 and 1979 assessments.

The panels selected objectives in four areas (two in math, two in reading):

1. **Math computation skills**: The focus was on basic arithmetic – addition, subtraction, multiplication and division.

   Sample item:

   $$\begin{array}{r} 321 \\ \times 12 \\ \hline \end{array}$$

   A. 3,852
   B. 852
   C. 963
   D. 3,752

2. **Math application skills**: The tasks presented in the test involved such problem solving skills as identifying the exact facts, translating written information into mathematical equations and recognizing when the information presented is inadequate for the task required. Such problems are often called "story problems." (It might interest the reader to know that the metric system was used in all of the items involving measurement.)

Sample item:

When Alaska became a state in 1959, it was the first new state in 47 years. Juneau, a city about 1600 kilometers (km) north of Seattle, Washington, became the capital of the new state. Until Alaska became a state, Arizona had been the last to join the United States. In what year did Arizona become a state?

39. What is one way to find the answer to this problem?

A. Add 1600 to 1959
B. Add 1600 to 47
C. Add 47 to 1959
D. Subtract 47 from 1959

3. **Word identification skills**: The objectives tested involved skills in phonics, word analysis and vocabulary building.

Sample item:

Read the sentences. Mark your answer sheet to show the word in the row which makes the second sentence mean *the opposite of* the first sentence.

40. The lake was very *deep*.

The lake was very _____.

A. long B. narrow C. shallow D. dark

4. **Reading comprehension skills**: The test required a student to read a paragraph or two and then respond to questions on that passage.

Sample item:

EASY COOKING DIRECTIONS

2 Servings	*4 Servings*
1-1/2 cups water	3 cups water
1/4 tsp. salt	1/2 tsp. salt
2/3 cup oatmeal	1-1/3 cups oatmeal

Bring water to boil. Add salt. Slowly stir in oatmeal. Stir gently over moderate heat for about 30 seconds. Do not cook longer than one minute. Remove from heat; serve immediately.

5. What do you do first when you fix the oatmeal?

A. Heat the water
B. Remove oatmeal from heat
C. Stir in oatmeal
D. Cook oatmeal for 30 seconds

A total of 47 objectives were tested in the fourth grade – 12 in mathematics computation, 5 in mathematics application, 15 in word identification skills, and 15 in reading comprehension skills. Except in mathematics application three test items were used to measure each objective, for a total of 180 items on the fourth grade test. A total of 48 objectives were tested in the eighth grade – 12 in mathematics computation, 8 in mathematics application, 13 in word identification skills and 15 in reading comprehension skills – for a total of 186 items. All items were pilot tested prior to use. Only the most effective items were retained; others were modified or replaced.

Readers may wonder why a standardized test already on the market was not used. Alaska citizens and educators who helped plan this assessment wanted to test objectives that reflected Alaska's particular goals and needs. Because no available test measured this complete set of objectives, an objective referenced test was developed.

This report does not compare individual students' scores to one another or to national norms. Instead, educators in local districts are expected to establish levels of satisfactory performance (known as *criterion levels)* against which to compare scores. The state averages reported herein should help educators set those criterion levels.

As a state we are more concerned about using assessment results to determine whether our students are attaining the learning objectives our curriculum specialists consider important than we are about comparing our students' performance to that of students in other states whose curriculum may or may not resemble our own.

RESULTS and INTERPRETATION

How Are the Results Interpreted?

You may be familiar with tests which compare students to a national average, or norm. Such tests, designed for use nationwide, are called norm referenced tests. In norm referenced tests, items are selected to insure variation among student scores.

The Alaska test is different. Each student's performance is compared to desired levels of performance on specific learning objectives. Students or classrooms are not compared to one another. The purpose of such testing is to find out whether any given student or group of students is learning the specific skills being measured. We do not select test items to insure variability among students.

How Well Did Our Students Do?

Because local school districts differ in the goals and objectives they emphasize in their schools, it is difficult to establish one criterion level that is appropriate for the whole state. Each school that participated in the assessment received its classroom results along with the statewide results on each objective. Thus school officials and teachers would know how their students had performed in relation to students statewide, but not in relation to students in other individual schools or districts.

Each district was asked to decide what level of performance was acceptable to them in each content area. They could set different standards (criterion levels) for each area. Thus it may not be appropriate, strictly on the basis of scores, to say that students did better in mathematics application than in mathematics computation. A district may set a higher criterion level for mathematics application than for mathematics computation.

The chart below shows the average number of items correctly answered by students statewide in the four content areas.

AVERAGE NUMBER OF ITEMS ANSWERED CORRECTLY BY STUDENTS STATEWIDE

	Mathematics				Reading			
	Computation		Application		Word Identification		Comprehension	
Year	4th (36)*	8th (36)	4th (54)	8th (66)	4th (45)	8th (39)	4th (45)	8th (45)
1977	26	22	31	37	32	29	26	27
1979	27	23	33	40	35	30	29	29

*Number in parenthesis indicates the total number of items on the subtest.

The slight increases in assessment results for 1979 were not sufficient to claim definite improvement in students' skills. Data must be collected over several years to confirm such a trend.

Did Results Vary According to Student and District Characteristics?

To better understand differences in performance among groups of students, three types of comparisons were made:

1. **District size and geographic area**
   Part of the sampling procedure involved separating districts on the basis of size and geography. The labels of large, medium, small and rural were applied as follows:

   Large – Enrollment over 12,000 students
   Medium – Enrollment over 1000 but less than 12,000
   Small – Enrollment of less than 1000 in incorporated cities
   Rural – The 21 "new" districts; Regional Education Attendance Areas, unorganized boroughs

   The percentage of students tested within each of these categories is approximately 59, 22, 8 and 11 respectively.

2. **Bilingual**
   Approximately 4 percent of the students tested were judged by their teachers as being bilingual on the basis that a language other than English was spoken at home.

3. **Reading problems**
   Teachers judged whether students had a mild or severe reading problem, or no reading problem. Generally, the percentage of students within each of these three categories was 14, 6, and 80 respectively.

NUMBER OF STUDENTS TESTED

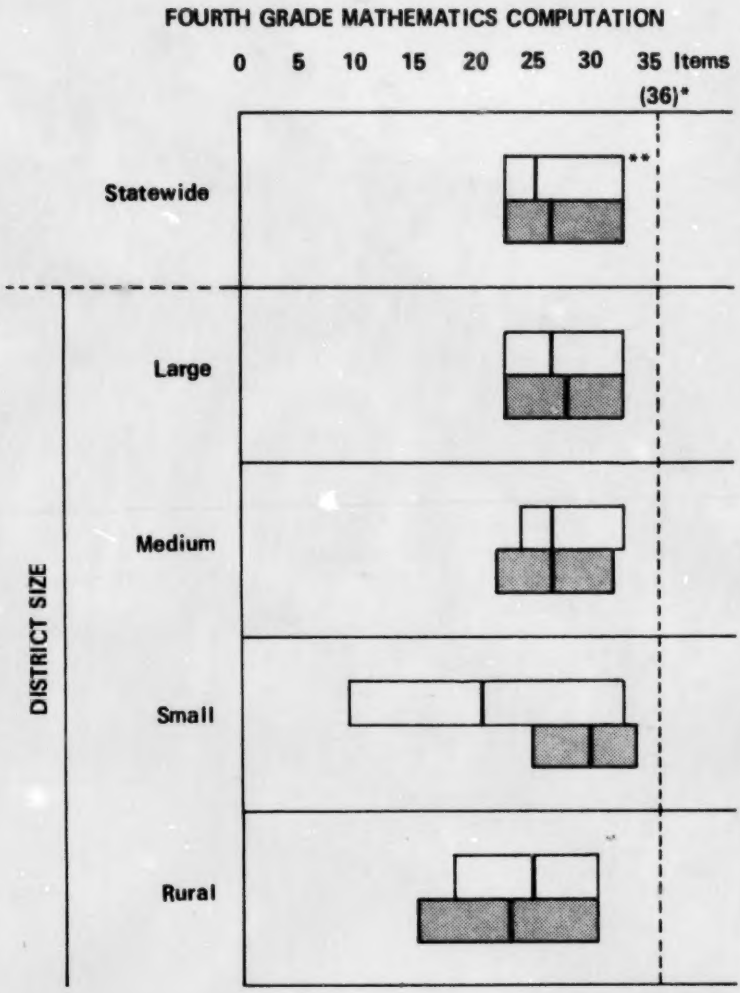
		Statewide	District Size			
			Large	Medium	Small	Rural
GRADE 4	1977	1197	666	270	100	161
	1979	1439	912	338	87	102

		Bilingualism		Reading Problems		
		Non-bilingual	Bilingual	None	Mild	Severe
GRADE 4	1977	1126	30	853	217	96
	1979	1091	52	866	181	88

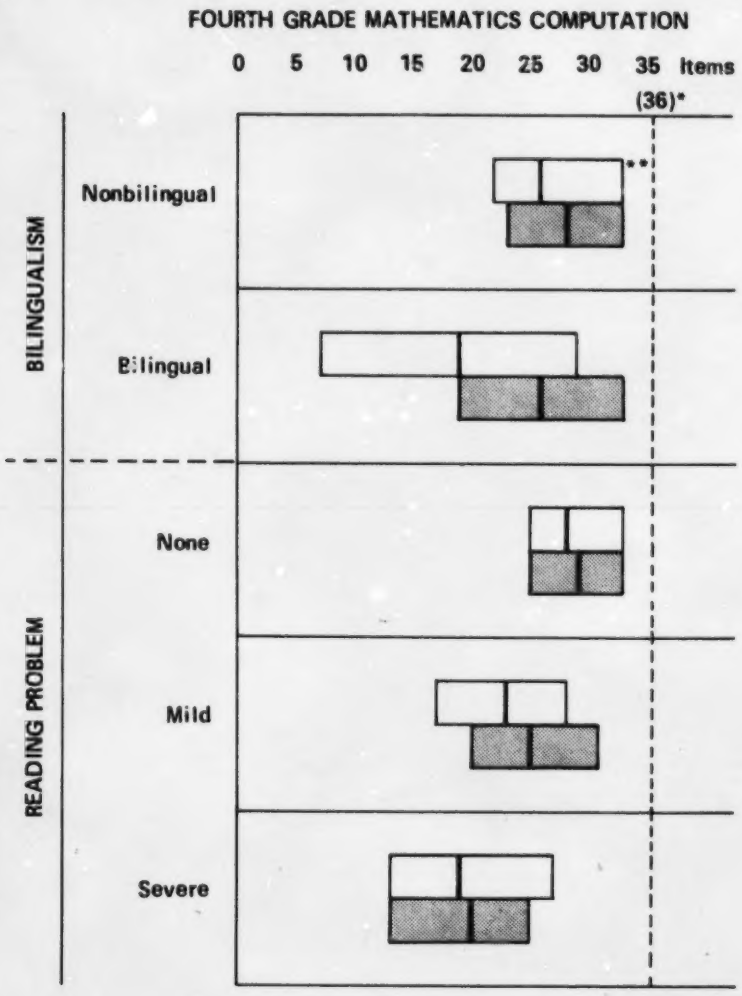
		Statewide	District Size			
			Large	Medium	Small	Rural
GRADE 8	1977	1243	660	287	111	185
	1979	1440	867	267	148	158

		Bilingualism		Reading Problems		
		Non-bilingual	Bilingual	None	Mild	Severe
GRADE 8	1977	1171	47	998	146	84
	1979	966	59	859	95	56

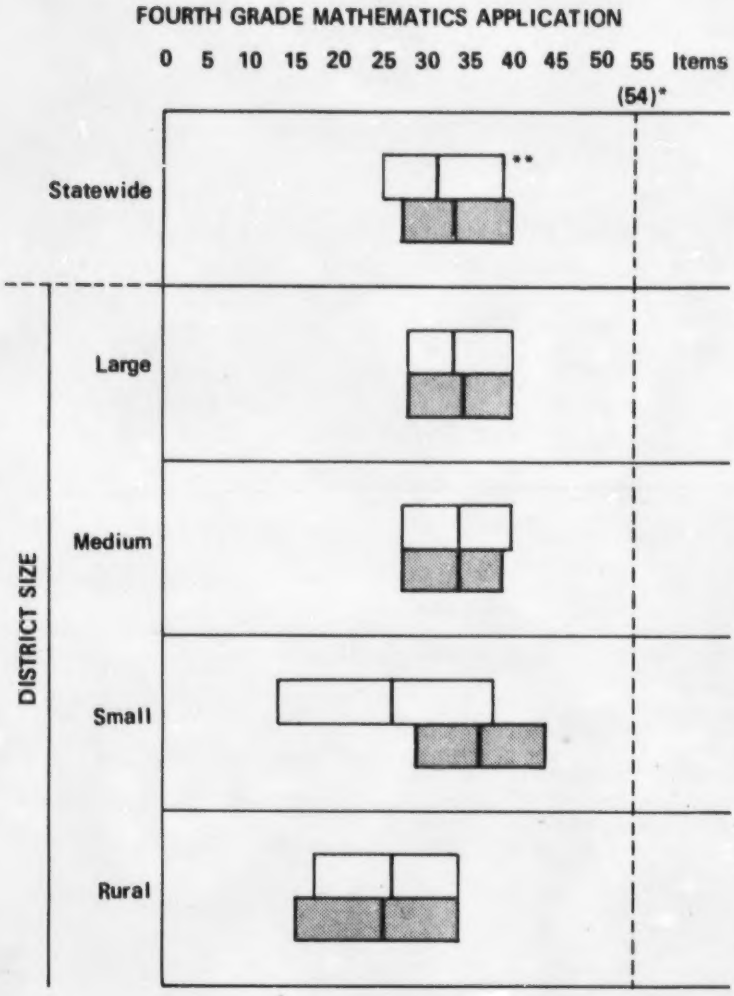
The figures on the next few pages display differences in performance according to the categories defined here. The bars on each chart indicate the range of scores for the middle 50 percent of students, The dark line within the bar indicates the average score for the group. For example, in the first figure, the first two bars represent the 1977 and 1979 statewide results for fourth graders in mathematics computation. The middle 50 percent of the students answered from 23 to 33 out of 36 items correctly in both 1977 and 1979. In other words, 25 percent of the students answered fewer than 23 items correctly and 25 percent answered more than 33 correctly. The average number of items answered correctly was 26 in 1977 and 27 in 1979.



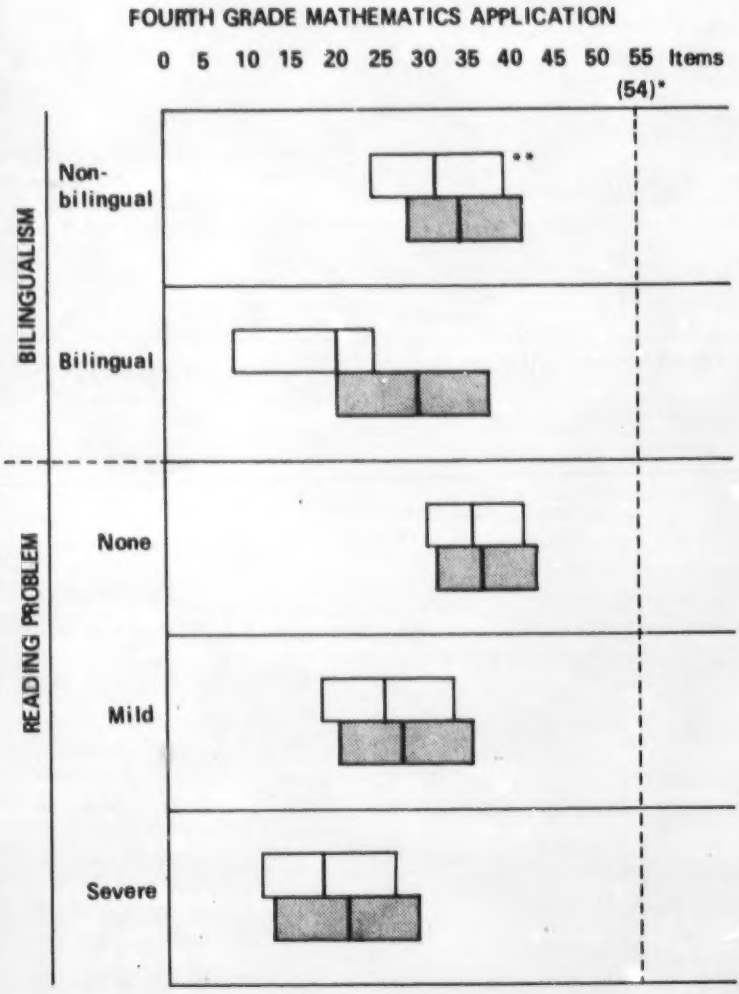
* Total number of items
** unshaded bars – 1977
shaded bars – 1979



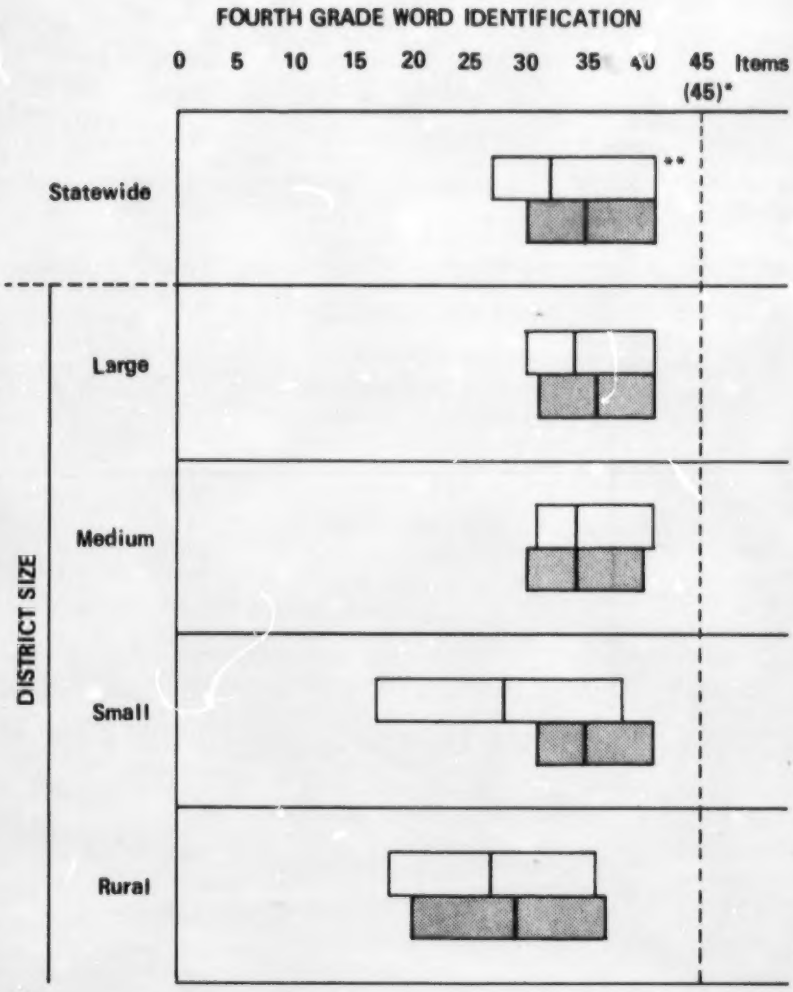
* Total number of items
** unshaded bars – 1977
shaded bars – 1979



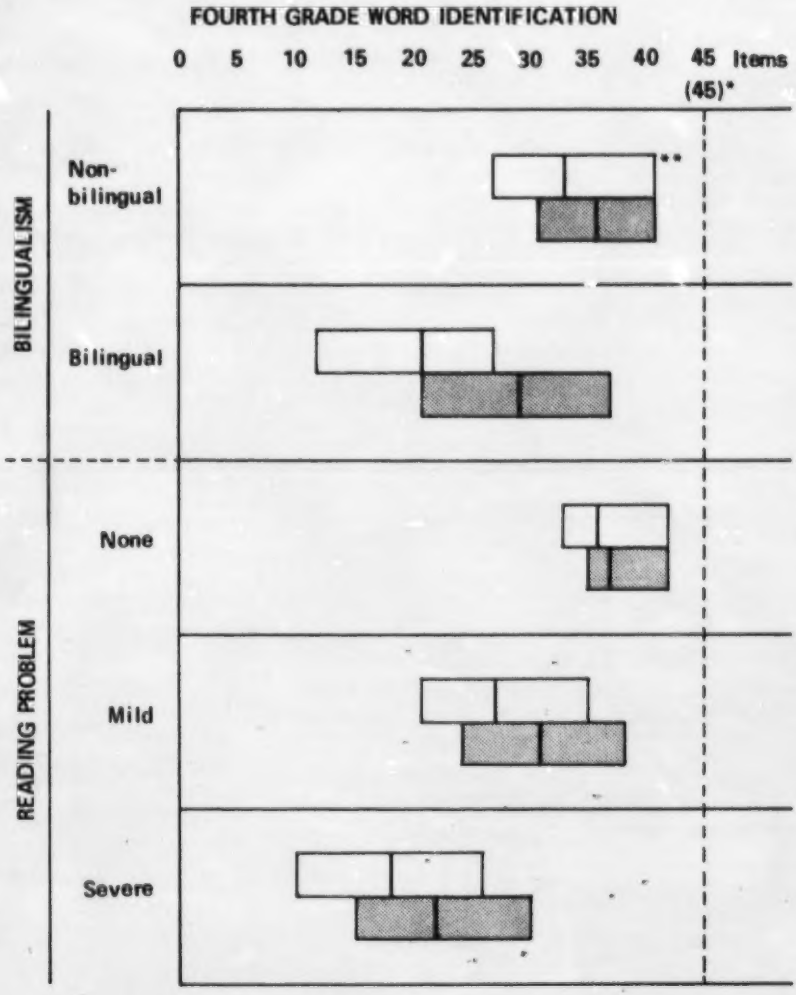
* Total number of items
** unshaded bars – 1977
shaded bars – 1979



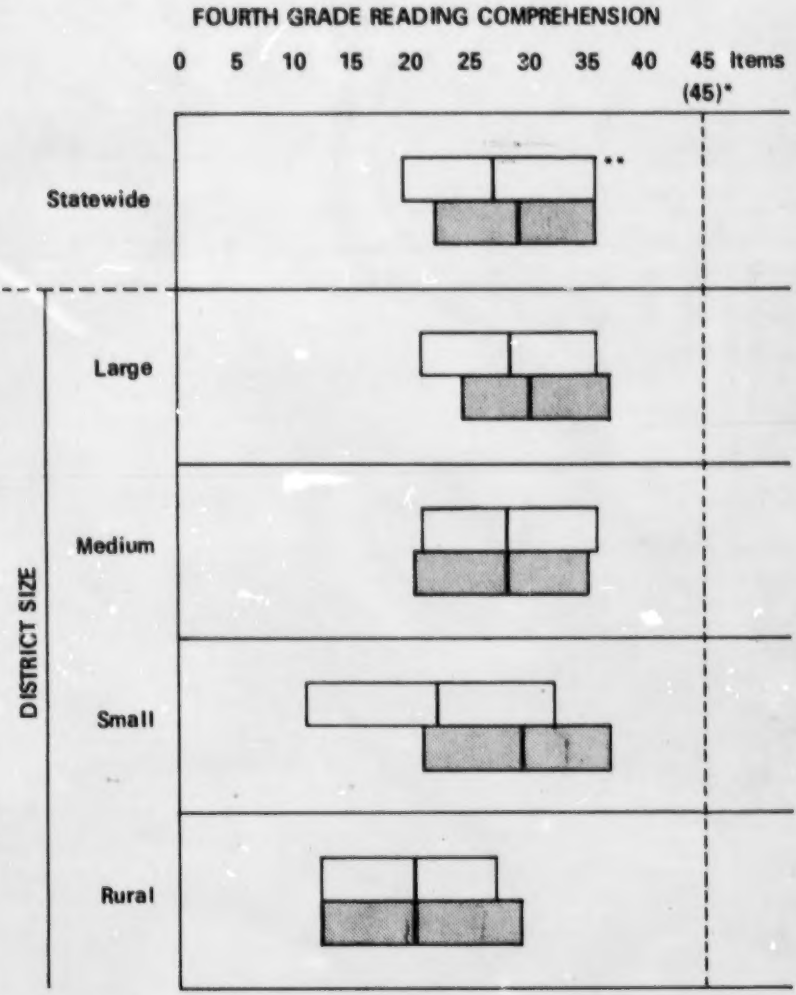
* Total number of items
** unshaded bars – 1977
shaded bars – 1979



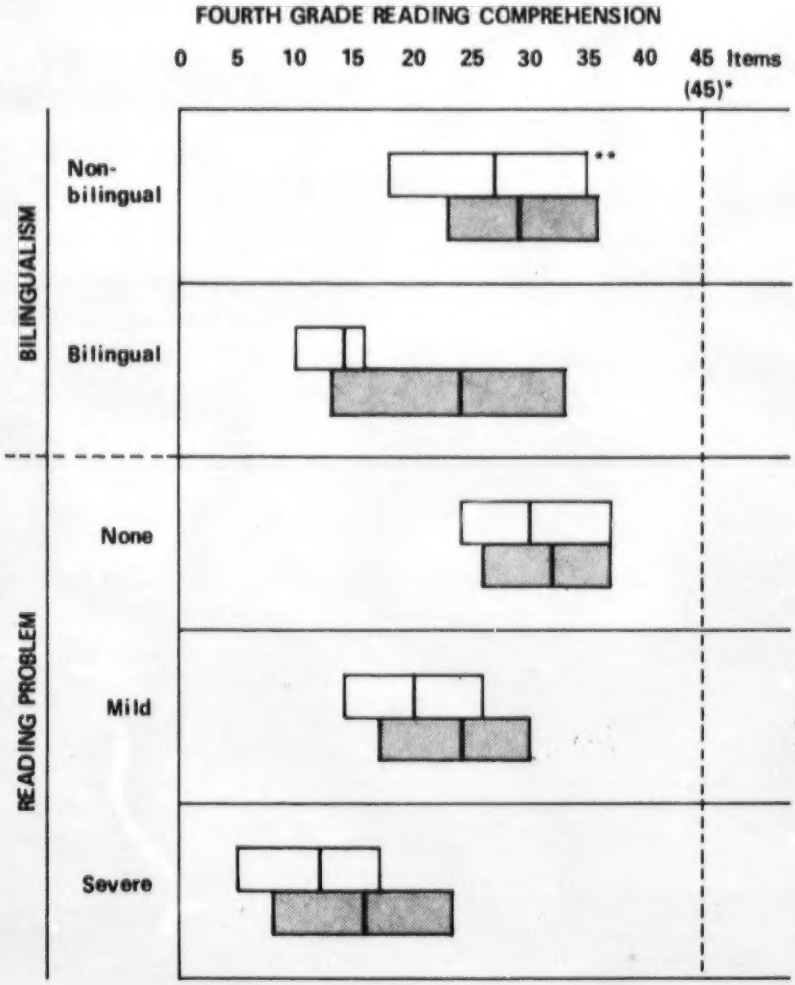
* Total number of items
** unshaded bars – 1977
shaded bars – 1979



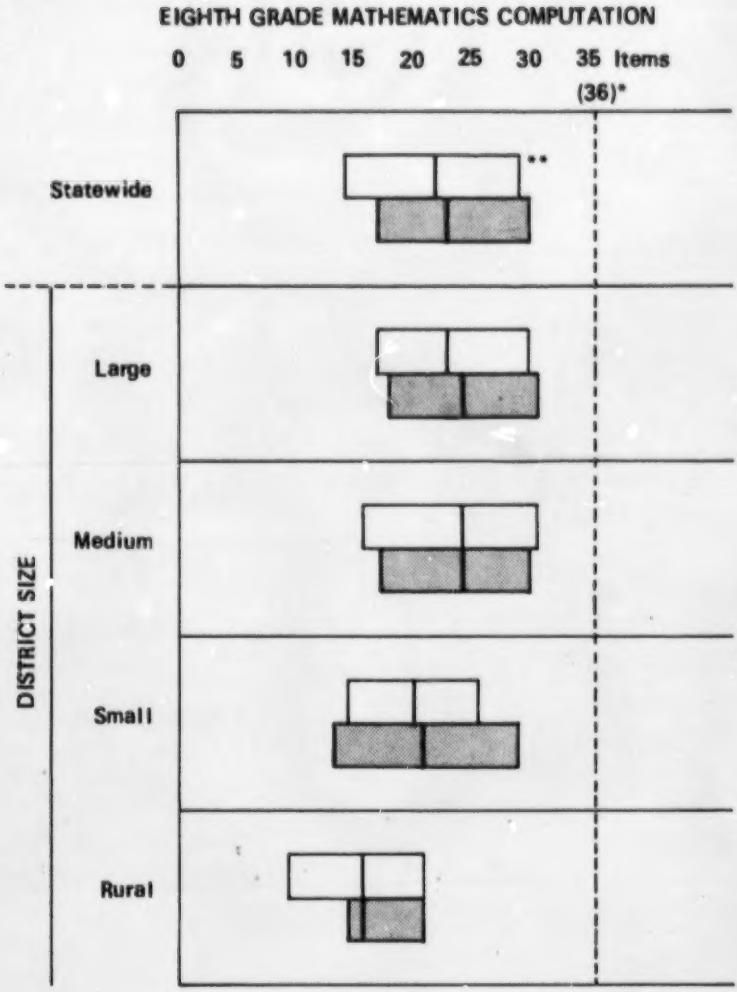
* Total number of items
** unshaded bars – 1977
shaded bars – 1979



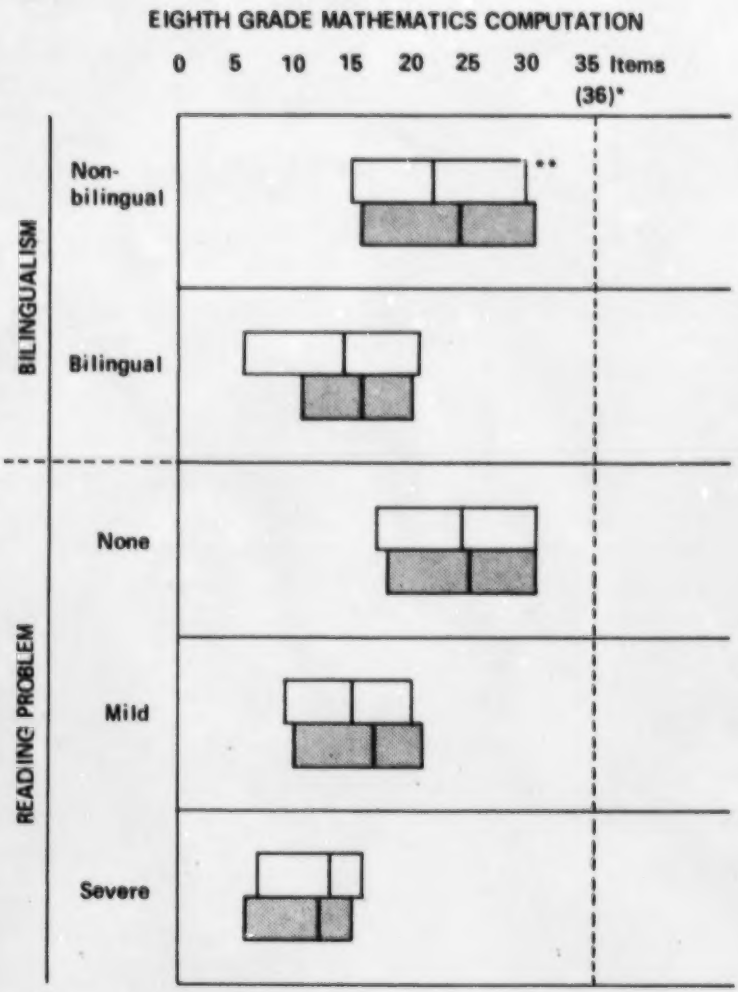
* Total number of items
** unshaded bars – 1977
shaded bars – 1979



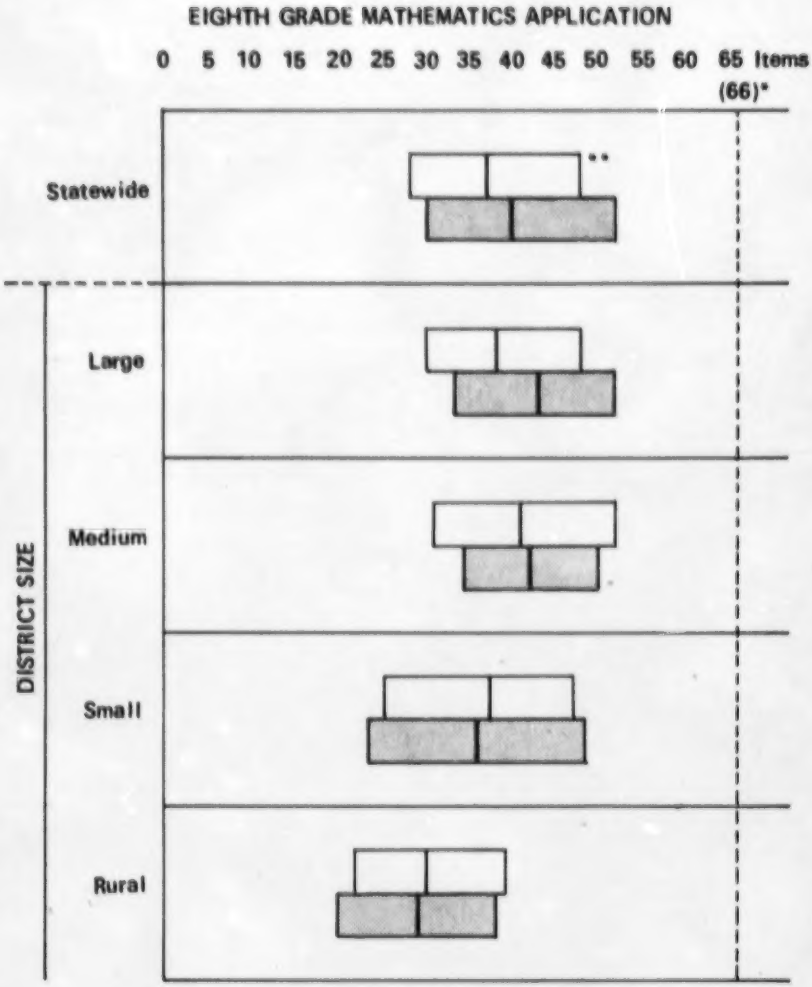
* Total number of items
** unshaded bars – 1977
shaded bars – 1979



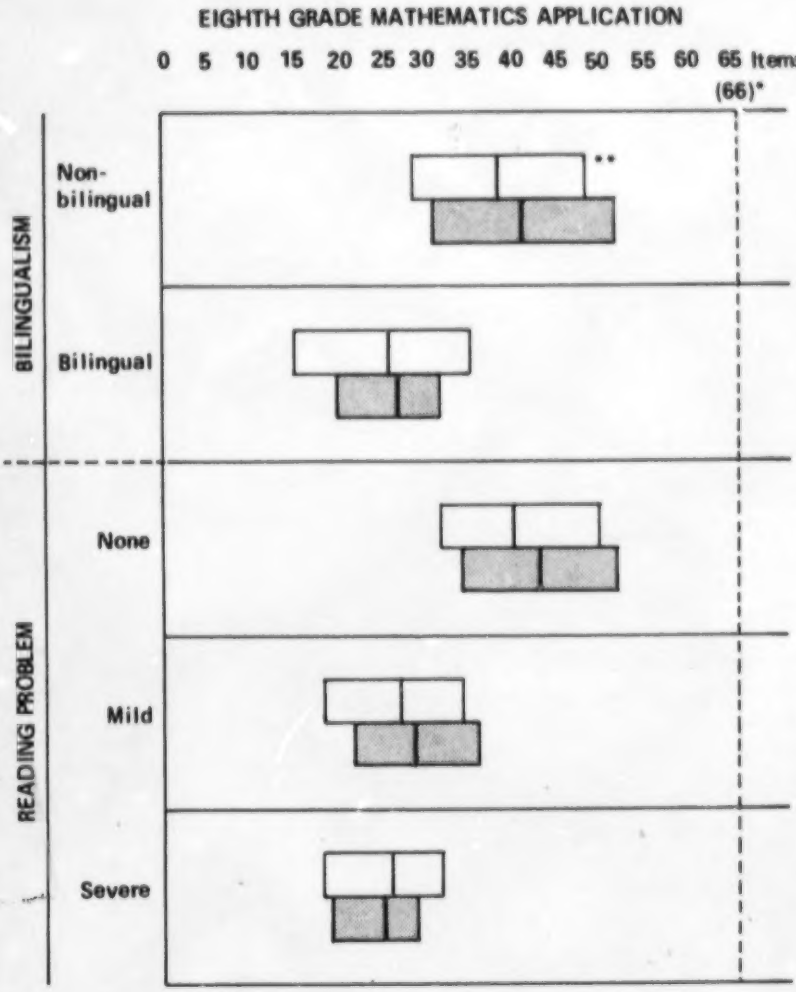
* Total number of items
** unshaded bars – 1977
shaded bars – 1979



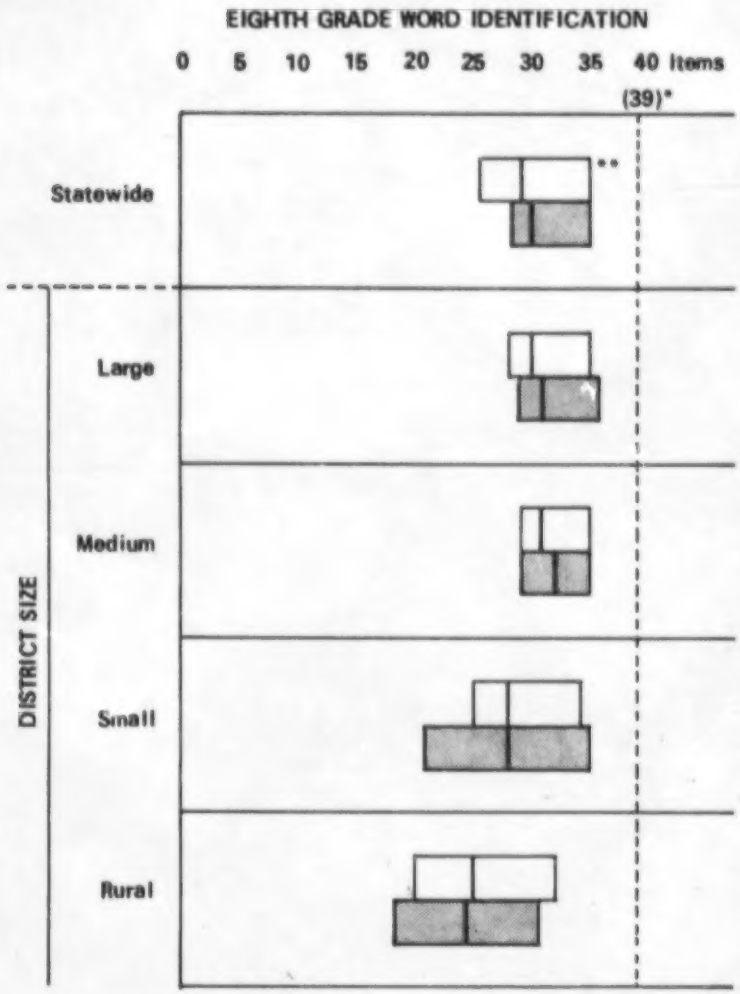
* Total number of items
** unshaded bars – 1977
shaded bars – 1979



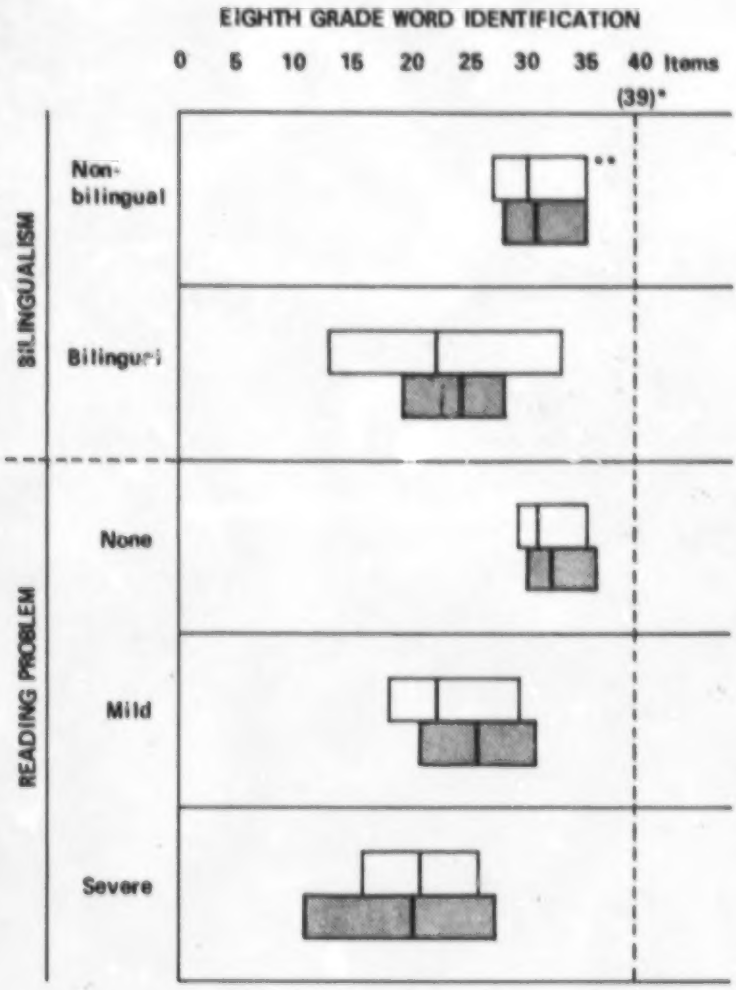
* Total number of items
** unshaded bars – 1977
shaded bars – 1979



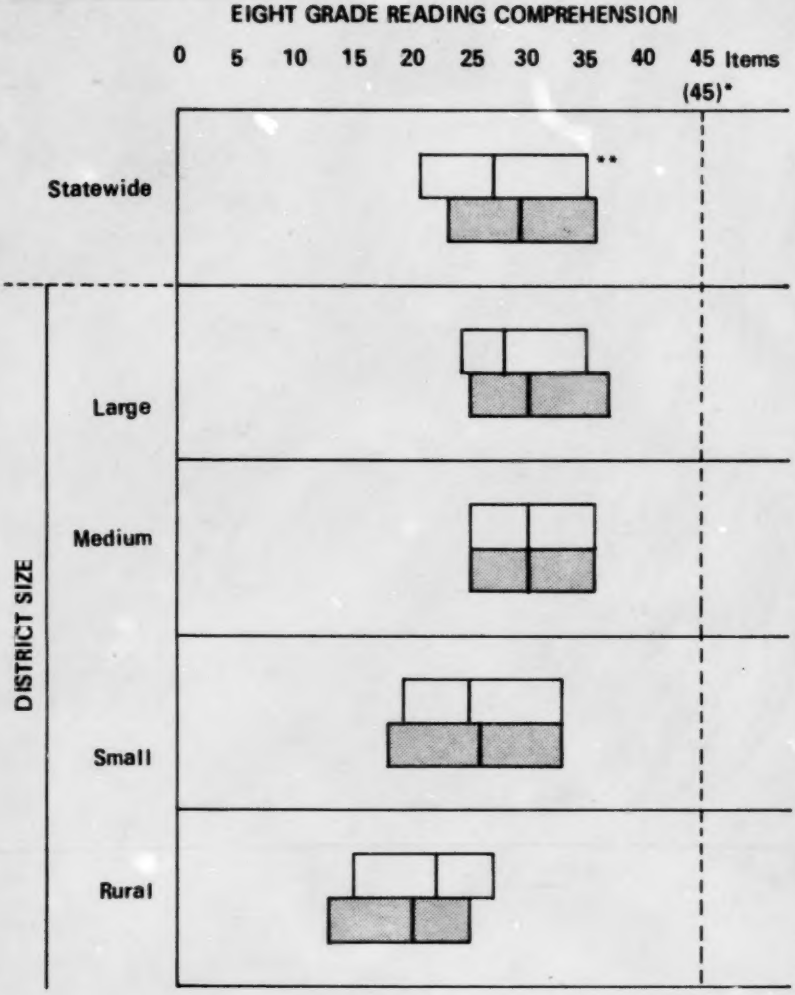
* Total number of items
** unshaded bars – 1977
shaded bars – 1979



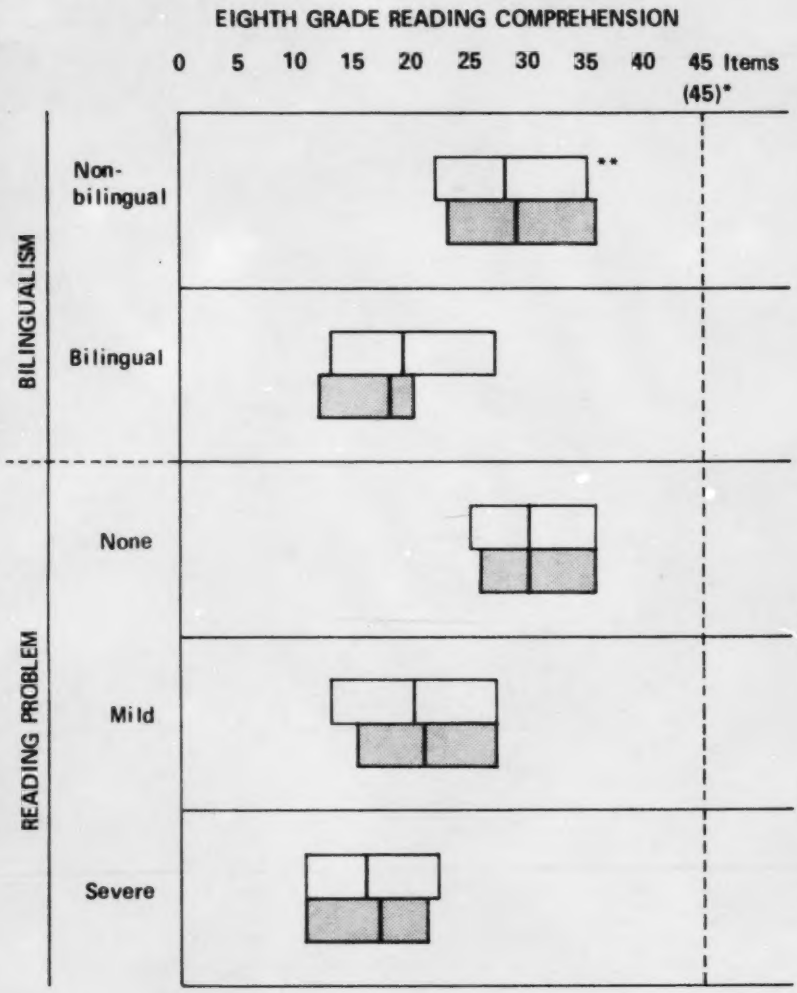
* Total number of items
** unshaded bars – 1977
shaded bars – 1979



* Total number of items
** unshaded bars – 1977
shaded bars – 1979



* Total number of items
** unshaded bars – 1977
shaded bars – 1979



* Total number of items
** unshaded bars – 1977
shaded bars – 1979

SUMMARY OF FOURTH GRADE RESULTS

Statewide, scores did not change substantially between 1977 and 1979 for fourth graders. Across all four content areas in 1977, students in small and rural districts scored lower than those in large or medium sized districts.

Bilingual students did noticeably better in 1979 than in 1977. However, since the number of bilingual students tested was small, these changes must be interpreted with great caution.

In both 1977 and 1979 students with reading problems did not do as well as those without such problems.

It is important to remember that causal statements cannot be made from these data alone. Rather the findings suggest relationships in need of further investigation. For example, why does district size seem to affect scores?

SUMMARY OF EIGHTH GRADE RESULTS

The charts for eighth graders show similar patterns across all four content areas. In each area students in large and medium sized districts performed better than those in small and rural districts for both 1977 and 1979. Non bilingual students performed better than bilingual students. As students' reading problems increased, their scores decreased. Statewide, the 1979 sample of students performed slightly better than the 1977 sample.

What Additional Information is Available?

Information about class performance has been provided to all schools that were selected for testing. In addition, schools that asked to participate, even though they were not selected as part of the sample, also received information on individual students.

Teachers can obtain additional useful information about individual students by using the Alaska Instructional Diagnostic System (AIDS) Information about AIDS and technical information on the testing system can be obtained from the Alaska Department of Education's Office of Planning and Research.